Saccone, N. L., Baurley, J. W., Bergen, A. W., David, S. P., Elliott, H. R., Foreman, M. G., Kaprio, J., Piasecki, T. M., Relton, C. L., Zawertailo, L., Bierut, L. J., Tyndale, R. F., & Chen, L. S. (2018). The value of biosamples in smoking cessation trials: a review of genetic, metabolomic, and epigenetic findings. *Nicotine and Tobacco Research*, *20*(4), 403-413. [ntx096]. https://doi.org/10.1093/ntr/ntx096

Peer reviewed version

Link to published version (if available):
10.1093/ntr/ntx096

Title:
The Value of Biosamples in Smoking Cessation Trials: A Review of Genetic, Metabolomic, and Epigenetic Findings

Author list:
Nancy L. Saccone, PhD[1]*, James W. Baurley, PhD[2], Andrew W. Bergen, PhD[2], Sean P. David, MD, DPhil[3], Hannah R. Elliott, PhD[4], Marilyn G. Foreman, MD, MS[5], Jaakko Kaprio, MD, PhD[6], Thomas M. Piasecki, PhD[7], Caroline L. Relton, PhD[4], Laurie Zawertailo, PhD[8], Laura J. Bierut, MD[9], Rachel F. Tyndale, PhD[10], Li-Shiun Chen, MD, MPH, ScD[9] on behalf of the Genetics and Treatment Networks of the Society for Research on Nicotine and Tobacco (SRNT)

[1] Department of Genetics, Washington University School of Medicine, St. Louis, Missouri, USA
[2] BioRealm, Culver City, California, USA
[3] Department of Medicine, Stanford University, Stanford, California, USA
[4] MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Bristol, UK.
[5] Pulmonary and Critical Care Medicine, Morehouse School of Medicine, Atlanta, Georgia, USA
[6] Institute for Molecular Medicine, University of Helsinki, Helsinki, Finland
[7] Department of Psychological Sciences, University of Missouri, Columbia, Missouri, USA
[8] Nicotine Dependence Service, Centre for Addiction and Mental Health, and Department of Pharmacology and Toxicology, University of Toronto, Toronto, Canada
[9] Siteman Cancer Center, Institute of Public Health, and Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri, USA
[10] Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, and Departments of Pharmacology & Toxicology and Psychiatry, University of Toronto, Toronto, Canada

*Correspondence to:
Nancy L. Saccone, Ph.D.
Department of Genetics
Washington University School of Medicine
St. Louis, MO 63110
USA
Tel: (314) 747-3263
E-mail: nlims@genetics.wustl.edu

Abstract

Human genetic research has succeeded in definitively identifying multiple genetic variants associated with risk for nicotine dependence and heavy smoking. To build on these advances and aid in reducing the prevalence of smoking and its consequent health harms, the next frontier is to identify genetic predictors of successful smoking cessation and also of the efficacy of smoking cessation treatments ("pharmacogenomics"). More broadly, additional biomarkers that can be quantified from biosamples also promise to aid "Precision Medicine" and the personalization of treatment, both pharmacological and behavioral. To motivate ongoing and future efforts, here we review several compelling genetic and biomarker findings related to smoking cessation and treatment. These include genetic variants in the nicotinic receptor subunit gene *CHRNA5*, variants in the nicotine metabolism gene *CYP2A6*, and the nicotine metabolite ratio. We also summarize reports of epigenetic changes related to smoking behavior. The results to date demonstrate the value and utility of data generated from biosamples in clinical treatment trial settings. This paper cross-references a companion paper in this issue that provides practical guidance on how to incorporate biosample collection into a planned clinical trial and discusses avenues for harmonizing data and fostering consortium-based, collaborative research on the pharmacogenomics of smoking cessation.

Implications

Evidence is emerging that certain genotypes and biomarkers are associated with smoking cessation success and efficacy of smoking cessation treatments. We review key findings that open potential avenues for personalizing smoking cessation treatment according to an individual's genetic or metabolic profile. These results provide important incentive for smoking cessation researchers to collect biosamples and perform genotyping in research studies and clinical trials.

1. Introduction

Cigarette smoking remains a leading preventable cause of global morbidity and mortality [1-2].Addressing this problem requires a two-pronged strategy both to prevent individuals from initiating cigarette use and to assist current smokers in the process of quitting. Despite substantial educational and policy-related initiatives to discourage smoking, extrapolation from recent trends suggests there are likely to be more than a billion smokers worldwide by 2025 [3]. Thus, for the foreseeable future, smoking cessation treatments will remain an indispensable component of tobacco control efforts. Although a number of effective smoking cessation interventions have been developed and disseminated, clinical trials of evidence-based treatments typically find that 60% or more of treated smokers relapse within 5-6 months [4]. Research that advances our understanding of the basic neurobiological determinants of smoking and relapse has potential to accelerate progress in smoking cessation treatment.

Recent characterization of the human genome, combined with advances in genotyping technology, have led to an explosion of new knowledge concerning the genetics of complex diseases [5]. Nicotine and tobacco research has made substantial contributions to this emerging literature, producing some of the largest replicable effects of genotypes on disease risk or related traits in neuropsychiatric genetics [6-9]. One exciting implication of this work is the possibility of new opportunities for precision medicine for smoking cessation [10-11], for which multiple treatments are available, both pharmacological and behavioral. Existing studies indicate that various biomarkers, including genotypes, predict responses to smoking cessation pharmacotherapies [12-13]. Importantly, genotypes are also associated with naturalistic cessation and the course of smoking-related disease, highlighting their clinical impact [14]. Ideally, as pharmacogenomic evidence accumulates and as the costs of genetic assays decrease, it will

become possible to tailor treatment strategies to individuals in routine clinical practice, improving treatment efficacy and minimizing adverse effects. Basic research into the genetics of smoking may also reveal new insights into the pathophysiology of nicotine dependence, identifying novel treatment targets.

Even with these successes, substantial challenges must be overcome to realize the clinical promise of this swiftly growing body of genetic research. One clear lesson from existing studies is that the effect sizes of individual common genetic variants on common, complex phenotypes are small in absolute magnitude, often accounting for ~1% of phenotypic variance in measures such as cigarettes per day or tobacco dependence (e.g. [15-18]). Detecting such small, but biologically meaningful, effects is made more complicated by the need to apply stringent statistical corrections for multiple testing when many variants are examined simultaneously. Owing to concerns about possible false positives and publication bias, genetic research places a premium on independent replication and meta-analysis [19-21]. Finally, assessment of rare variants may be necessary to fully characterize genetic contributions to smoking and nicotine dependence [15, 22-25].

These considerations indicate that the way forward requires the use of very large samples, probably numbering in the tens of thousands of individuals, especially if the phenotypes are based on self-report. Such large-scale studies rely on collaboration across many independent investigators and have been very productive in human genetics generally [26]. Another strategy to increase statistical power is to use biomarkers [27] that are relevant for smoking, such as cotinine [28-30], exhaled carbon monoxide [18], or the nicotine metabolite ratio [31-36]. For these biomarkers, the proportion of variance accounted for by genetic loci is several fold greater than for standard smoking measures, and association studies find multiple genome-wide significant findings with

much smaller sample sizes. Epigenetic studies also have the potential to leverage modest sample sizes, as large effect sizes have been reported for the association between levels of DNA methylation (a commonly studied epigenetic process) and smoking-related phenotypes. Ultimately, to establish and extend the clinical utility of genetic research for smoking cessation, genetic variants and other biomarkers must be assayed and analyzed together with clinically relevant and detailed phenotypic data from the same individuals. Important traits include relapse status, time to various behavioral and clinical milestones (including comorbidities and consequent disease), adherence to and adverse effects from pharmacotherapies, and potential mediators of treatment effects (e.g., withdrawal symptoms, craving, lapse antecedents and consequences).

Clearly, much can be gained by moving beyond a traditional model of clinical research relying on small teams of investigators working independently. Progress will require a concerted, "big science" effort. This article is one of a pair of review papers produced through collaboration between the Genetics Network and the Treatment Network of the Society for Research on Nicotine and Tobacco (SRNT). The overarching aim of these reviews is to facilitate this big science approach by encouraging clinical researchers to routinely collect and bank biospecimens, even when they have no immediate plans to analyze the samples or link biomarkers to clinical outcomes. To the extent this becomes a standard practice, nicotine and tobacco researchers will create a reservoir of genetically informative clinical data that can be tapped to speed the pace of scientific discovery and foster the development of precision medicine approaches. Future collaborative analyses of this collective resource will be expedited if investigators work together and communicate to determine scientific priorities and promote consistent phenotypic assessments to address them.

To set the stage for this kind of effort, the current article reviews selected examples of how data generated from biosamples - specifically genotypes, metabolite measurements, and epigenetic profiles - have been valuable for tobacco research, including evidence concerning the moderation of treatment effects by genotype. The first sections present the two most well-established genetic findings for smoking behavior to date, the *CHRNA5* cholinergic nicotinic receptor subunit gene and the *CYP2A6* nicotine metabolism gene. An additional section presents epigenetic findings relevant for smoking and cessation research. A companion article [37] offers practical information on how to incorporate biosample collection into a planned clinical trial, considers issues surrounding data curation and sharing, and looks forward to consortium-based approaches for pooling and analyzing the data. Table 1 introduces a glossary of genetics and genomics terms that will be used throughout this review paper and also the companion article [37]. Table 2 summarizes key genetic findings for *CHRNA5* and *CYP2A6* that this review will highlight.

2. Genetic variants in *CHRNA5*, the α5 cholinergic nicotinic receptor subunit gene

CHRNA5 and smoking:

Genetic variants in *CHRNA5*, the gene encoding the α5 cholinergic nicotinic receptor subunit, have come to prominence for their associations with smoking traits including nicotine dependence [38, 56], smoking quantity and heaviness, e.g. cigarettes per day [16, 40-41, 43], and biomarkers of smoking such as cotinine [28-30], total nicotine equivalents [57], and carbon monoxide [18]. In particular, studies of European-ancestry smokers have identified and replicated the non-synonymous single nucleotide polymorphism (SNP) rs16969968, which causes an amino acid change from aspartic acid to asparagine that alters *in vitro* function of *CHRNA5* [58]. Furthermore, this variant and its correlated proxies are associated with smoking-related diseases such as lung

cancer [38, 45, 59] and chronic obstructive pulmonary disease [46-47]. This *CHRNA5* region also harbors multiple independently associated loci; for example, a distinct group of variants, including rs880395 and rs588765, are associated with smoking when rs16969968 is included as a covariate, and are also associated with *CHRNA5* mRNA levels (gene expression) in brain [60-62]. Other variants in the surrounding genes (e.g. *IREB2*, and additional nicotinic receptor subunit genes *CHRNA3* and *CHRNB4*) are highly correlated (in linkage disequilibrium) with these *CHRNA5* variants and thus also show associations with smoking.

Less work has been performed in non European-ancestry populations. Nevertheless, *CHRNA5* variants are associated with smoking behaviors in large studies of African-ancestry and Asian-ancestry smokers [42, 63]. A genome-wide meta-analysis of 32,389 African-ancestry participants, including 15,547 ever-smokers, observed genome-wide significant association between cigarettes per day and rs2036527, a SNP in the distal enhancer region of *CHRNA5* [42]. Though this SNP does not confer an amino acid change, in African-ancestry populations it tags the *CHRNA5*-expression-associated haplotype [60-62]. Another meta-analysis targeted *CHRNA5* variants in smokers of African (N=10,912), Asian (N=6,889), and European (N=14,786) ancestry [63]; all three ancestry groups showed consistent association of the non-synonymous SNP rs16969968 with heavy smoking, despite considerable allele frequency differences, ranging from 42% in European-ancestry samples to 7% in African Americans and 3% in Asians. Other non-European populations with differing haplotype structure or low frequency for rs16969968 may show association with the *CHRNA5/A3/B4* region though not necessarily this variant [51, 57, 64].

These genetic associations with nicotine dependence and related smoking behaviors have motivated several studies of the utility of *CHRNA5* variants to predict efficacy of smoking cessation treatment in randomized clinical trials. The results, outlined below, demonstrate both

the potential of genetic information to enable personalized treatment, and also the importance of careful interpretation that accounts for study design and potential environmental effects.

CHRNA5 and smoking cessation treatment

Evidence from multiple treatment trials indicates that the *CHRNA5* variants involved in nicotine dependence risk can distinguish between smokers who are likely to respond strongly to pharmacologic treatment and those who are unlikely to benefit [49-50, 65-68].

We focus here on evidence that the functional variant rs16969968 is associated with treatment response. In 1,075 European-American smokers in a randomized, placebo-controlled clinical trial, genetic analyses compared high- and low-risk haplotypes for nicotine dependence, defined by two SNPs: rs16969968 (the amino acid change D398N) and rs880395 (associated with *CHRNA5* mRNA expression changes) [49]. Biochemically confirmed 7-day point prevalent abstinence at end-of-treatment was analyzed as the cessation outcome. All participants received behavioral counseling and were randomized to placebo or active treatment (nicotine replacement therapy (NRT), buproprion, and combination of both). Results showed a significant interaction between treatment and haplotype ($\chi 2 = 8.97$, df=2, omnibus p = 0.02). Specifically, individuals with the high-risk haplotype had the lowest abstinence rate when treated with placebo plus counseling; however, adding pharmacologic treatment for smoking cessation significantly increased the likelihood of smoking cessation back to levels comparable to the low-risk haplotype group. In contrast, those with the low-risk haplotype received little effect from pharmacologic treatment for smoking cessation; they quit just as successfully with placebo plus counseling. These results translate to a dramatic difference between the two genetic profiles in the number needed to treat (NNT), the number of people who must receive treatment for one person to benefit. The NNT was 4 for smokers with the high-risk profile, versus over 1000 for

smokers with the low-risk profile [69-70]. Moreover, the high-risk haplotype is quite common, with a frequency of 21% in this European-American sample, translating to meaningful impact on a population level.

The interaction between *CHRNA5* genetic status and pharmacological treatment was confirmed in a combined analysis of 2,633 European-ancestry smokers from eight randomized clinical trials [50], including data from the trial in Chen et al. [49]. As before, *CHRNA5* SNPs representing the amino acid change and the mRNA expression change were highlighted. In particular, at the available proxy SNPs selected for analysis, the minor alleles for the amino acid change and the expression change were significantly associated with reduced six month 7-day point prevalent abstinence in the placebo group, and with increased abstinence in participants randomized to nicotine replacement therapy. The minor allele for the amino acid change was also associated with significantly reduced abstinence in the placebo group at end of treatment.

In non European-ancestry smokers, there is also evidence that *CHRNA5* is associated with treatment response. Multiple candidate SNPs in the *CHRNA5-A3-B4* region were analyzed for association with two randomized, placebo controlled trials of African-Americans for smoking cessation [66]. The functional SNP rs16969968 had low allele frequency and did not show consistent association; however the A allele of rs2036527, which in African-ancestry cohorts is associated with cigarettes per day [42], was associated with lower abstinence with active pharmacotherapy during treatment and at end of treatment for nicotine gum, and when analyzing pooled data from the nicotine gum and bupropion trials [66].

However, some studies of *CHRNA5* in treatment trials have been less conclusive. We describe these next, and discuss how these varied reports highlight the importance of considering study design and potential environmental factors when interpreting results.

A recent meta-analysis of four clinical trials found no evidence that rs16969968 or its proxy rs1051730 increased smoking cessation in smokers receiving nicotine replacement therapy [71]. Although this result may at first appear to argue against a role for *CHRNA5* in treatment response, in fact it is consistent with the prior evidence that *CHRNA5* variants *interact* with treatment to significantly affect cessation. To see such an interaction, it is necessary to analyze both the non-pharmacologically treated (placebo) group and the treated group: carriers of the risk variants quit poorly on placebo, but those treated by NRT were "rescued" to attain cessation rates comparable to the low-risk genetic group. Thus a comparison of genetic status only in pharmacologically treated smokers, as in [71], is not designed to inform about this interaction.

In a separate, placebo-controlled clinical trial of 654 European-ancestry smokers, *CHRNA5* genotype by treatment interactions were specifically analyzed and found to be not statistically significant [72]. This suggests that the effect may not be robust over different trials or populations of smokers. Still, the direction of effect trended in the expected direction, with lower cessation rates in placebo-treated, genetically high-risk smokers, but no statistical evidence of rescue by NRT.

The varied results highlight the need to improve statistical power via increased sample sizes, and to include understudied racial groups and account for diverse population histories. One challenge is that individual SNP effects discovered in large samples may be modest; however such findings are still valuable because they implicate specific genes and biological pathways and can improve polygenic score estimates. Also important is the consideration of non-genetic (environmental) factors that are well-established predictors of smoking cessation, as these may be involved in differences seen across trials. The strong *CHRNA5* effect reported in [49] was observed in a population of smokers who displayed strong motivation to quit. Similarly, a

naturalistic study of cessation in pregnant women smokers, who are generally expected to have strong motivation to quit, found significant association between *CHRNA5* risk genotype and reduced abstinence [48]. It is plausible that a setting of strong motivation or pressure to quit may highlight contrasts between genetically high-risk and low-risk smokers, thus more clearly manifesting that genetically high-risk smokers have reduced ability to quit without pharmacologic treatment. It is this reduced ability that underlies the significant genotype-by-treatment interaction in [49] and [50]. Other important factors may include partner smoking [73], treatment adherence [74], and adverse effects [75]. Overall, assessment of motivation and other modifying factors in clinical trials will be valuable to fully characterize the role of genetics in treatment response and cessation.

In summary, significant *CHRNA5* effects on treatment efficacy have been observed in several, though not all, studies. Genetic analyses of additional, well-phenotyped treatment trials that include placebo treatment arms will be crucial to establishing whether this promising biological predictor ultimately will have clinical value for smokers seeking to quit. The current evidence provides strong motivation for further biosample collection in clinical trials for genetic research. Future studies in larger samples, as well as collaboration across studies, are important to increase power to examine specific medications for potential differences in effect.

3. *CYP2A6* and nicotine metabolism biomarkers

Nicotine is the main addictive component in tobacco and to a large extent responsible for the maintenance of smoking, and individuals vary in their nicotine clearance rates [76-77]. Thus biomarkers of nicotine metabolism - including genotypes at nicotine metabolism genes and direct measures of metabolite levels in blood, plasma, or urine - are valuable for the study of smoking

behavior. Unlike cotinine and carbon monoxide, which are used as biomarkers of smoking behavior such as amount smoked, and which demonstrate associations with *CHRNA5* [18, 28-30], biomarkers of nicotine metabolism focus on representing variability in metabolism, which contributes in specific ways to smoking behavior.

CYP2A6 and smoking

The *CYP2A6* gene on chromosome 19 encodes the enzyme involved in metabolizing nicotine and also its major metabolite, cotinine [78-79]. *CYP2A6* variants, and neighboring variants in *EGLN2*, a gene associated with response to hypoxia [80], are among the consistent genome-wide significant signals associated with cigarette consumption [16, 40, 51]. *EGLN2* variation has recently been shown to be associated with cigarette consumption independently of *CYP2A6* variation, and also associated with nicotine dependence measures and with smoking efficiency (the ratio of carbon monoxide over cigarette consumption), though not with nicotine metabolism [81].

CYP2A6 and biomarkers of CYP2A6 activity

CYP2A6 enzyme activity can be represented by the ratio of plasma or salivary *trans*-3'-hydroxycotinine/cotinine or urinary *trans*-3'-hydroxycotinine plus *trans*-3'-hydroxycotinine glucuronide/cotinine (3HC/COT), a biomarker known as the nicotine metabolite ratio (NMR) [31]. Higher NMR values indicate faster nicotine metabolism.

NMR is influenced by individual genetic variation and also other factors beyond *CYP2A6* genotype, such as hormones. From twin modeling, the proportion of variation of NMR attributable to additive genetics is estimated to be ~67-81% and 47% for NMR measured from plasma and urine, respectively [32, 82], and is reduced to ~49% and 42% upon adjustment for genotyped *CYP2A6* variants [82]. At the same time, the influences of environmental and demographic variables explain only 8% of the variance in NMR [83]. These estimates suggest that

NMR is influenced by these *CYP2A6* alleles and also by additional heritable variation. Indeed, a candidate gene study of drug metabolizing enzyme and transporter (DMET) gene variation confirmed a major role for *CYP2A6* variation in nicotine metabolism, attaining genome-wide significant association of specific *CYP2A6* variants with NMR [33]. Furthermore, recent genome-wide studies of NMR have identified genome-wide significant associations with *CYP2A6* variants [32, 34, 52]. In particular, all three studies identified a *CYP2A6* intervening sequence four (IVS4) variant, rs56113850, that is associated with reduced NMR in European ancestry individuals. The first of these GWAS meta-analyses used cotinine-verified current smokers (N = 1,518) from multiple Finnish cohorts [32]; methylation analyses were also conducted, and are described in the next section on epigenetics. Of the hundreds of genome-wide significant signals discovered in the region of chromosome 19q13.2, the most significant was the *CYP2A6* IVS4 variant rs56113850 (p=5.8E-86). Conditional analyses identified three additional independent genome-wide significant associations in the region. The top three SNPs account for 31% of NMR variance in these Finnish cohorts. The IVS4 SNP was later replicated in a separate GWAS meta-analysis of the laboratory-based NMR in 49 African, 51 Asian, and 212 European ancestry individuals [34]. The most statistically significant results were a SNP proximal to *CYP2A6* (rs12459249, p = 2.5E-18), and the IVS4 SNP (rs56113850, p = 6.6E-18); another region in the *CYP2A6-CYP2A7* intergenic region yielded two signals with p <3E-10. Bioinformatic analyses of rs56113850 further connected this variant to reduced transcription in liver and lung and reduction of incorporation of exon 5 of *CYP2A6* in liver [34]. Further replication of rs56113850 was observed in a third GWAS of 2,239 current smokers of African American, Native Hawaiian, White, Latino, and Japanese American ethnicity; this IVS4 SNP was the most

significant signal in the multi-ethnic GWAS and the only SNP that was genome-wide significant in each of the five ethnic groups with sample sizes ranging from 311 to 674 participants.

An alternative biomarker of CYP2A6 activity directly uses genotypes at variants at *CYP2A6* in a regression model to predict the ratio of cotinine over the sum of nicotine and cotinine [COT/(NIC+COT)]. This genotype-based activity metric accounted for 73% of the variance in COT/(NIC+COT) 30 minutes after oral administration of labeled compound, and 44% of the variance in NMR at six hours after oral nicotine administration, in European-American subjects [84]. The regression model derived from those data provides a biomarker of CYP2A6 activity that can be calculated from seven *CYP2A6* variants that are reported to be relatively common in Europeans, and does not require direct nicotine metabolite measurements.

This genotype-based metric was subsequently studied for its associations with smoking behaviors in a distinct sample of smokers. In both European-American and African-American smokers, the metric was associated with cigarettes per day [53].

Biomarkers of CYP2A6 activity and smoking cessation treatment:

We have outlined how CYP2A6 activity can be represented by the metabolite-based NMR and by an alternative metric based on *CYP2A6* genotypes. Both these biomarkers have shown associations with efficacy of smoking cessation treatment in clinical treatment trials, indicating their potential as tools to personalize treatment.

Several studies have evaluated NMR and efficacy of nicotine replacement therapy (NRT) and found that low NMR (indicating slower metabolism) corresponds to greater nicotine patch effectiveness. First, an open-label, randomized, intention to treat strategy compared transdermal nicotine vs. nicotine nasal spray in 480 smokers who all received behavioral therapy [85]. The hypothesis was that higher NMR would be associated with lower cessation success rates with

NRT, measured via the primary outcome of prolonged abstinence at the end of treatment and at 6 months' follow-up. The effect of pre-treatment NMR on abstinence was tested in the total sample and in each treatment group controlling for pre-treatment nicotine dependence (Fagerström score), Body Mass Index, race, and sex. NMR was associated with the effectiveness of transdermal nicotine: among smokers treated with transdermal nicotine, low NMR corresponded to greater abstinence. Additionally, higher NMR predicted lower nicotine concentrations and more severe cigarette cravings after one week of treatment. There was no association of NMR and nicotine nasal spray, which the authors speculated could be related to pharmacodynamic differences. This correspondence of low NMR to greater nicotine patch effectiveness has been replicated in other clinical trials [54, 86] and in a community-based trial [87]. In African-American smokers, an overall association of low NMR with greater quit success was observed in both the NRT arm and the placebo arm [88].

The association of NMR with smoking cessation was furthermore evaluated in a prospective, biomarker-stratified clinical trial comparing transdermal nicotine vs. varenicline [35]. In this 11-week, multicenter, randomized, placebo-controlled, double-blinded, intention to treat trial, motivated smokers were assigned to placebo (placebo pill/placebo patch), NRT (active patch/placebo pill), or varenicline (active pill/placebo patch). All subjects also received behavioral counseling. The primary endpoint was biochemically verified 7-day point prevalence abstinence at the end of treatment (EOT). The primary hypothesis was an NMR-by-treatment interaction at EOT. Of the 1246 participants, 662 were characterized as slow metabolizers of nicotine (NMR <0.31) and 584 were considered normal/rapid nicotine metabolizers (NMR > 0.31). At EOT and at six months, varenicline was significantly more effective than transdermal nicotine in normal (p-values of 0.001 and 0.03) but not in slow (p-values of 0.56 and 0.51)

metabolizers. In a longitudinal model incorporating EOT, 6-month, and 12-month time points, the NMR-by-treatment interaction was significant [ratio of odds ratios = 1.96 (95% CI 1.11 – 3.46) p = 0.02]. In slow metabolizers compared to normal metabolizers, an NMR-by-treatment interaction found greater side effect severity from varenicline vs. placebo [β = -1.06 (95% CI -2.08 to -0.03), p =0.044]. Thus for normal metabolizers, varenicline was found to be superior to transdermal nicotine, corresponding to an NNT in this group of 4.9 for varenicline versus 26.0 for nicotine patch. Transdermal nicotine was equivalent to varenicline in slow metabolizers, and slow metabolizers were significantly more likely to report side effects from varenicline compared to placebo.

The genotype-based metric developed in [53, 84] is an alternative biomarker that shows association with treatment efficacy. This metric was used to classify 709 European-ancestry participants as slow (lowest quartile of the metric) versus normal (the remaining quartiles) metabolizers, in a randomized trial of placebo, active (nicotine replacement therapy, bupropion) and combined active therapies [89 55]. Among smokers receiving placebo, slow metabolizers were less likely to relapse at end of treatment compared to normal metabolizers (hazard ratio (HR)=0.40, p=0.013). In addition, Chen et al identified significant effects of NRT in reducing smoking relapse among fast metabolizers (HR=0.50, p=4.4E-6), but not among normal metabolizers (HR=0.93, p=0.75). In contrast, there were no significant differences in treatment effect by metabolism metric among those receiving bupropion. The differences in relapse rate by pharmacotherapy and genotype-based nicotine metabolism metric generated a significant interaction in the entire sample (HR=2.55, p=0.016). For those randomized to NRT versus placebo, the number needed to treat (NNT) differed significantly by nicotine metabolism status (2.9 for normal metabolizers and >1,000 for slow metabolizers).

Together, the findings outlined above indicate that further work to develop, compare, and test biomarkers of nicotine metabolism has the potential to pave the way for personalized smoking cessation treatment. One priority - and opportunity - is to more comprehensively address how nicotine metabolism speed impacts the efficacy of NRT, before clinical translation. As reviewed above, multiple clinical trials that used NMR to define metabolism speed found that slow metabolizers respond to NRT, with greater benefit than normal/fast metabolizers [35, 54, 85-86]. In contrast, a clinical trial that used a *CYP2A6* genotype-based metric to characterize nicotine metabolism speed found that slow metabolizers do not benefit from NRT compared to placebo, while normal/fast metabolizers benefit greatly [55]. These contrasting results may be due to a combination of several factors, including differences in the trial populations and in the way metabolism speed was measured. Reconciling these results should be possible with further biospecimen collection and genotyping in clinical trials that will allow comparison of these metabolite-based and genotype-based measures in the same clinical trial subjects. Finally, large trials involving non-European-ancestry smokers are needed: in African-Americans, analyses of *CHRNA5* [66] and nicotine metabolism [88] indicate that both influence cessation success, but further work is needed to understand the direction of effect and the role of pharmacologic treatment.

4. Epigenetics

Epigenetic processes define non-genetic modifications to the DNA that can alter gene regulation. The most commonly studied epigenetic process is DNA methylation, in which methyl groups attach to Cytosine-Guanine (CpG) dinucleotides in the DNA sequence with subsequent impact on gene expression [90]. With the advent of genome-wide arrays for DNA methylation measurement, it is possible to measure DNA methylation at single CpG sites throughout the

genome, analogously to genome-wide SNP arrays [91]. For the majority of population-based studies to date, DNA from peripheral blood cells has been the sample source used to measure DNA methylation, as other tissues are less readily accessible in large samples. The widespread use of this method has enabled replication efforts of findings in the field. Here we describe DNA methylation findings relevant to smoking. In particular, methylation levels at certain loci have been shown to be responsive to exposure to tobacco smoke, suggesting potential for application as markers of smoking status, including successful cessation.

Summary of epigenome-wide association studies for smoking:

Current tobacco smoking is robustly associated with DNA methylation in blood at a distinct set of loci across the genome. These findings have been comprehensively replicated in a number of population studies, using arrays [92-96] and also other laboratory technologies [92, 95-96]. A recent systematic review indicates that smoking-related methylation signatures may provide informative biomarkers not only for current smoking but also life-time exposure [97]. Of almost 200 replicated methylation sites associated with smoking exposure to date, the most well-known locus is *AHRR* (Aryl-Hydrocarbon Receptor Repressor). A CpG site annotated as cg05575921 on the Illumina Infinium HumanMethylation450 BeadChip (to date the most commonly used platform for this type of analysis, interrogating over 485,000 methylation sites), located in an intronic region of AHRR, is the site most commonly associated with smoking in both adults [96, 98-101] and adolescents [102]. The effect size of smoking on methylation at this unique site is so large (>20% difference in DNA methylation between exposed and unexposed) that it is detectable even in low powered studies of small sample size. Methylation profiles have also been shown to be disrupted in children exposed to maternal smoking *in utero* [93, 103-104]. Furthermore, methylation measures at this site have been used to identify nascent smokers [94].

The majority of studies conducted to date have been in European-ancestry populations but methylation differences appear to be very similar in South Asian and African American populations [99-100, 105]. As well as being a biomarker of smoking exposure, DNA methylation has also been shown to predict disease outcomes: *F2RL3* methylation is strongly associated with mortality [106], and *AHRR* methylation may be associated with subclinical atherosclerosis [107].

One of the major challenges in methylation studies is difficulty in inferring causal relationships. On a genome-wide scale, smoking-associated methylation changes can either be consequences of smoking, or predispose to smoking behavior. Further complications are caused by the dynamic nature of DNA methylation. Methylation at a given CpG site may reverse quickly after smoking cessation, others may reverse slowly, while some may be irreversible.

DNA methylation as a tool to assess smoking status and history:

Reliable assessment of smoking behavior and smoking history can be challenging, especially in higher risk settings where accurate quantification is most needed and where self-report is often least reliable. Examples of such groups include pregnant women who under report smoking during pregnancy when they know they are expected to reduce or quit tobacco use [108] or adolescents [109]. Cotinine or exhaled carbon monoxide levels can be used to detect recent smoking but are generally only able to detect smoking exposure within hours or a few days [110-113]. The concept of using DNA methylation as an archive of historical exposure has been mooted [114] and has been used to distinguish between never, former, and current smokers in recent studies [100, 115].

DNA methylation and NMR:

A recent genome-wide association study of NMR in three Finnish cohorts reported 719 significantly associated SNPs on chromosome 19q13, surrounding *CYP2A6* [32]. The authors postulated that other gene regulatory mechanisms, including DNA methylation, may influence NMR. They interrogated the genome-wide significant SNPs to test whether they are associated with DNA methylation levels, i.e. are methylation quantitative trait loci (meQTLs). Methylation values of 16 CpG sites on 19q13 were found to be associated with genotypes at the genome-wide significant SNPs. Further, causal inference testing demonstrated that methylation levels at one CpG site tagged by cg08551532 (in *DLL3*, ~1.3 Mbp proximal to *CYP2A6*) mediated the effect of some of the SNPs on NMR.

DNA methylation as a tool to evaluate smoking cessation:

In addition to being associated with smoking *per se*, several CpG sites, including those at *F2RL3*, *GPR15*, *LRRN3* and *AHRR*, are sensitive to cumulative smoking exposure and to cessation, with methylation levels reverting towards unexposed levels over time [95-96, 116]. A cross-sectional study of over 3500 individuals reported that *F2RL3* methylation decreased with smoking up to a maximum exposure of approximately 40 pack years. In former smokers, methylation was shown to increase with time since cessation until approximately 20-25 years after quitting, at which time methylation levels in former and never smoker were almost indistinguishable [116]. *AHRR* findings are similar, with methylation levels in smokers gradually approaching levels of never smokers over a period of approximately 25 years [96]. At other loci, the reversion of methylation patterns following cessation appears to be more rapid with some loci being indistinguishable between former and never smokers within 3 months [101].

Although there is undoubtedly still work to be done to better define and characterize DNA methylation patterns in response to smoking behavior and cessation, this feature is of

particular interest to those interested in smoking cessation strategies. In this context DNA methylation may provide a useful tool to measure treatment efficacy, for example by using this biological measure rather than self-report which may be subject to bias. Because smoking cessation trials measure outcomes in weeks or months, such methylation changes would need to take place on a comparable time scale to be informative.

Given the sensitive and specific response of the peripheral blood methylation signature to exposure to tobacco smoke, there is great potential for application in the context of smoking cessation strategies. Two immediate examples of how this information could enhance existing studies are providing a refined measure of smoking behaviors (duration and heaviness of smoking) and predicting historical smoking behavior in the absence of self-reported data. Given the substantial effect sizes seen in DNA methylation in response to tobacco smoke, it is likely that this measure may be useful even in small sample sizes. DNA methylation patterns can vary between tissue and cell types and while it is certainly clear that blood methylation patterns reflect smoking behaviors, there is good evidence that patterns are similar in other accessible samples including buccal cells sampled from saliva [117]. Furthermore, cheaper site specific methods are being developed meaning this type of measurement will be readily accessible for those considering using epigenetics to measure smoking exposure and monitor smoking cessation.

5. Summary/Conclusions

This review has highlighted the value of collecting and generating data from biosamples from clinical smoking cessation treatment trials. First, evidence is accumulating that smokers' genotypes measured from DNA have utility for predicting not only risk of nicotine dependence and heavy smoking, but also cessation success under different treatments. The necessary DNA

can be extracted from blood, buccal swabs, or saliva for example, as discussed further in the companion paper Chen et al. Nicotine metabolism can be represented by genotypes and also by metabolite measurements in blood, plasma, or urine; such biomarkers of metabolism are proving informative for smoking cessation and treatment outcomes. Epigenetic studies in smokers, especially DNA methylation as a marker of smoking and cessation, constitute another exciting area of current investigation that is enabled by biosample collection.

Our review of genetic findings focused on the two most well-established genes that have been associated with nicotine dependence and smoking behaviors: *CHRNA5* and *CYP2A6*. Both these genes also show strong evidence for involvement in response to pharmacologic smoking cessation treatment. Our goal here was to highlight the strongest results in the expectation that eventually, multiple findings with strong evidence will be available to inform personalized treatment. Genomics implementation is a growing area of research focused on optimizing the integration of genomic findings into community and clinical settings; such research will ensure that basic science findings on established genetic variants (e.g., *CHRNA5*) are used to motivate smoking cessation and personalize treatment.

Limitations of the results reviewed here also highlight emerging opportunities. For example, many of these findings need to be investigated in other populations beyond those of European ancestry. Also, single SNPs discovered by GWAS tend to have modest effects on the smoking traits typically available in large consortium collaborations; this is due in part to the polygenic nature of behavioral traits and also the noise inherent in coarse measures such as cigarettes per day. Therefore it is important to move beyond individual SNP effect sizes, for example using polygenic risk scores. Also, other study designs that utilize more refined phenotypes can complement large efforts based on broader, more readily assessed phenotypes.

Ultimately, clinical use of genomic testing to personalize smoking cessation treatment will also require establishing both clinical validity and also clinical utility, with development of laboratory diagnostic tests approved for specific contexts of care [118]. These needs signal opportunities for researchers in smoking cessation and in genetics to work together on new genetic and epigenetic investigations. Much of the prior progress in identifying and validating genetic effects on nicotine dependence and smoking quantity came from collaborative consortium efforts. Similarly, as stated in the Introduction, we can tackle the genetics of smoking cessation and treatment through concerted, collaborative efforts. While literature-based meta-analyses and systematic reviews offer one way to combine evidence across multiple studies [119], they are subject to heterogeneity across studies due to differences in ascertainment, phenotype definitions, and the statistical models tested and reported. Greater power and accuracy is achieved through *de novo,* collaborative meta-analyses in which the primary investigators work together to reduce heterogeneity by harmonizing phenotypes and applying consistent statistical models to re-analyze data, prior to meta-analysis. This collaborative approach has been successful for smoking quantity [16, 40-43, 63] and has begun to be applied for cessation and treatment [50]. Data can be further optimized when studies communicate early in the study design phase and agree to use standardized phenotypic measures, such as those described in the PhenX Toolkit [120]. Collaborations among experts from many domains - clinical, pharmacological, epidemiological, genetic, statistical, and more - will enable new insights into the biological underpinnings of smoking cessation treatment efficacy, and cessation success.

6. Funding.

We would like to acknowledge the following funding support for our authors: From the National Institute on Drug Abuse (NIDA): R01 DA026911 (NLS), DA030398 (LSC), R01 DA038076 (LSC), HHSN271201300004C (JWB), R43 DA041211 (JWB, AWB), R21DA033813 (AWB), R01 DA017441 (SPD), U54 MD010724 (SPD), R01DA036583 (LJB), DA020830 (RFT); Medical Research Council Integrative Epidemiology Unit at the University of Bristol (MC_UU_12013/2) (HRE, CLR), postdoctoral research fellowship award from the Oak Foundation (HRE); U54MD008149 from the NIMHD (MGF); grants 265240 & 263278 from the Academy of Finland (JK); Sigrid Juselius Foundation (JK); Global Research Awards in Nicotine Dependence: WI 206396 and WS2391913 (Pfizer, Inc) (LZ); Canadian Cancer Society: 703187 and 703404 (LZ); Canadian Institutes of Health Research: DC0190SR (LZ); support of a Canada Research Chair in Pharmacogenomics (RFT).

7. Declaration of Interests.

JWB is an owner and employee of BioRealm and AWB is an employee of BioRealm, which offers commercial services related to the Smokescreen Genotyping Array. LJB is listed as an inventor on Issued U.S. Patent 8080371 "Markers for Addiction" covering the use of certain SNPs in determining the diagnosis, prognosis, and treatment of addiction, and served as a consultant for the pharmaceutical company Pfizer in 2008. The spouse of NLS is also listed as an inventor on the above patent. SPD has consulted for BaseHealth, Inc., which develops predictive platforms for population health management. JK has consulted for Pfizer. RT has consulted for the pharmaceutical company Apotex. The remaining authors declare no conflict of interest.

8. Acknowledgments.

This paper and a second companion paper were conceived, developed, and written by the Genetics and Treatment Workgroup of the Society for Research on Nicotine and Tobacco (SRNT): Drs. James W. Baurley, Andrew W. Bergen, Li-Shiun Chen, Sean P. David, Hannah R. Elliott, Marilyn G. Foreman, Jaakko Kaprio, Thomas M. Piasecki, Caroline L. Relton, Nancy L. Saccone, Laurie Zawertailo. We thank SRNT and especially the SRNT Genetics Network and the SRNT Treatment Network for their support and enthusiasm for this and other collaborative opportunities for genetics and treatment researchers. We thank Dr. Jennifer Ware for her input during early stages of this work, and Dr. Timothy Baker and Dr. Caryn Lerman for their assistance in reviewing an earlier draft. Thank you also to Mona Johnson for providing administrative and communications support.

We thank the SRNT Genetic Network and Treatment Network (Drs. Leonie Brose, Marissa Ehringer, Lisa Fucito, Anne Joseph, Anu Loukola, and Megan Piper) for review and comments.

9. Literature cited.

1. Jha P, Peto R. Global effects of smoking, of quitting, and of taxing tobacco. *N Engl J Med.* Jan 2 2014;370(1):60-68.

2. WHO. *WHO global report on mortality attributable to tobacco*. Geneva: World Health Organization; 2012.

3. Bilano V, Gilmour S, Moffiet T, et al. Global trends and projections for tobacco use, 1990-2025: an analysis of smoking indicators from the WHO Comprehensive Information Systems for Tobacco Control. *Lancet.* Mar 14 2015;385(9972):966-976.

4. Schlam TR, Baker TB. Interventions for tobacco smoking. *Annu Rev Clin Psychol.* 2013;9:675-702.

5. Green ED, Guyer MS. Charting a course for genomic medicine from base pairs to bedside. *Nature.* Feb 10 2011;470(7333):204-213.

6. Gelernter J. Genetics of complex traits in psychiatry. *Biol Psychiatry.* Jan 1 2015;77(1):36-42.

7. Kendler KS, Chen X, Dick D, et al. Recent advances in the genetic epidemiology and molecular genetics of substance use disorders. *Nat Neurosci.* Feb 2012;15(2):181-189.

8. Loukola A, Hallfors J, Korhonen T, Kaprio J. Genetics and smoking. *Curr Addict Rep.* Mar 1 2014;1(1):75-82.

9. Bierut LJ. Genetic vulnerability and susceptibility to substance dependence. *Neuron.* Feb 24 2011;69(4):618-627.

10. Bierut LJ, Johnson EO, Saccone NL. A glimpse into the future - Personalized medicine for smoking cessation. *Neuropharmacology.* Jan 2014;76 Pt B:592-599.

11. Chen LS, Horton A, Bierut L. Pathways to precision medicine in smoking cessation treatments. *Neurosci Lett.* May 18 2016.

12. Mamoun M, Bergen AW, Shieh J, Wiggins A, Brody AL. Biomarkers of response to smoking cessation pharmacotherapies: Progress to date. *CNS Drugs.* Apr 17 2015.

13. Bough KJ, Lerman C, Rose JE, et al. Biomarkers for smoking cessation. *Clin Pharmacol Ther.* Jun 2013;93(6):526-538.

14. Chen LS, Hung RJ, Baker T, et al. CHRNA5 risk variant predicts delayed smoking cessation and earlier lung cancer diagnosis-a meta-analysis. *J Natl Cancer Inst.* May 2015;107(5).

15. Olfson E, Saccone NL, Johnson EO, et al. Rare, low frequency and common coding variants in CHRNA5 and their contribution to nicotine dependence in European and African Americans. *Mol Psychiatry.* Aug 4 2015.

16. The Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet.* May 2010;42(5):441-447.

17. Vrieze SI, McGue M, Iacono WG. The interplay of genes and adolescent development in substance use disorders: leveraging findings from GWAS meta-analyses to test developmental hypotheses about nicotine consumption. *Hum Genet.* Jun 2012;131(6):791-801.

18. Bloom AJ, Hartz SM, Baker TB, et al. Beyond cigarettes per day. A genome-wide association study of the biomarker carbon monoxide. *Ann Am Thorac Soc.* Sep 2014;11(7):1003-1010.

19. Chanock SJ, Manolio T, Boehnke M, et al. Replicating genotype-phenotype associations. *Nature.* Jun 7 2007;447(7145):655-660.

20. Munafo MR, Flint J. Meta-analysis of genetic association studies. *Trends Genet.* Sep 2004;20(9):439-444.

21. Sullivan PF. Spurious genetic associations. *Biol Psychiatry.* May 15 2007;61(10):1121-1126.

22. Haller G, Druley T, Vallania FL, et al. Rare missense variants in CHRNB4 are associated with reduced risk of nicotine dependence. *Hum Mol Genet.* Feb 1 2012;21(3):647-655.

23. Xie P, Kranzler HR, Krauthammer M, et al. Rare nonsynonymous variants in alpha-4 nicotinic acetylcholine receptor gene protect against nicotine dependence. *Biol Psychiatry.* Sep 15 2011;70(6):528-536.

24. Wessel J, McDonald SM, Hinds DA, et al. Resequencing of Nicotinic Acetylcholine Receptor Genes and Association of Common and Rare Variants with the Fagerstrom Test for Nicotine Dependence. *Neuropsychopharmacology.* 2010;35(12):2392-2402.

25. Yang J, Wang S, Yang Z, et al. The contribution of rare and common variants in 30 genes to risk nicotine dependence. *Mol Psychiatry.* Nov 2015;20(11):1467-1478.

26. Evangelou E, Ioannidis JP. Meta-analysis methods for genome-wide association studies and beyond. *Nat Rev Genet.* Jun 2013;14(6):379-389.

27. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther.* Mar 2001;69(3):89-95.

28. Keskitalo K, Broms U, Heliovaara M, et al. Association of serum cotinine level with a cluster of three nicotinic acetylcholine receptor genes (CHRNA3/CHRNA5/CHRNB4) on chromosome 15. *Hum Mol Genet.* Oct 15 2009;18(20):4007-4012.

29. Munafo MR, Timofeeva MN, Morris RW, et al. Association between genetic variants on chromosome 15q25 locus and objective measures of tobacco exposure. *J Natl Cancer Inst.* May 16 2012;104(10):740-748.

30. Ware JJ, Chen X, Vink J, et al. Genome-Wide Meta-Analysis of Cotinine Levels in Cigarette Smokers Identifies Locus at 4q13.2. *Sci Rep.* 2016;6:20092.

31. Dempsey D, Tutka P, Jacob P, 3rd, et al. Nicotine metabolite ratio as an index of cytochrome P450 2A6 metabolic activity. *Clin Pharmacol Ther.* Jul 2004;76(1):64-72.

32. Loukola A, Buchwald J, Gupta R, et al. A Genome-Wide Association Study of a Biomarker of Nicotine Metabolism. *PLoS Genet.* Sep 2015;11(9):e1005498.

33. Bergen AW, Michel M, Nishita D, et al. Drug Metabolizing Enzyme and Transporter Gene Variation, Nicotine Metabolism, Prospective Abstinence, and Cigarette Consumption. *PLoS One.* 2015;10(7):e0126113.

34. Baurley JW, Edlund CK, Pardamean CI, et al. Genome-wide association of the laboratory-based nicotine metabolite ratio in three ancestries. *Nicotine Tob Res.* Apr 25 2016;18(9):1837-1844.

35. Lerman C, Schnoll RA, Hawk LW, Jr., et al. Use of the nicotine metabolite ratio as a genetically informed biomarker of response to nicotine patch or varenicline for smoking cessation: a randomised, double-blind placebo-controlled trial. *Lancet Respir Med.* Jan 9 2015;3(2):131-138.

36. Patel YM, Park SL, Han Y, et al. Novel Association of Genetic Markers Affecting CYP2A6 Activity and Lung Cancer Risk. *Cancer Res.* Oct 01 2016;76(19):5768-5776.

37. Chen L-S, Zawertailo L, Piasecki TM, et al. Leveraging genomic data in smoking cessation trials in the era of Precision Medicine: Why and how *Nicotine and Tobacco Research.* Companion paper submitted in parallel.

38. Thorgeirsson TE, Geller F, Sulem P, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature.* Apr 3 2008;452(7187):638-642.

39. Hancock DB, Reginsson GW, Gaddis NC, et al. Genome-wide meta-analysis reveals common splice site acceptor variant in CHRNA4 associated with nicotine dependence. *Transl Psychiatry.* 2015;5:e651.

40. Thorgeirsson TE, Gudbjartsson DF, Surakka I, et al. Sequence variants at CHRNB3-CHRNA6 and CYP2A6 affect smoking behavior. *Nat Genet.* May 2010;42(5):448-453.

41. Saccone NL, Culverhouse RC, Schwantes-An TH, et al. Multiple Independent Loci at Chromosome 15q25.1 Affect Smoking Quantity: a Meta-Analysis and Comparison with Lung Cancer and COPD. *PLoS Genet.* 2010;6(8):e1001053.

42. David SP, Hamidovic A, Chen GK, et al. Genome-wide meta-analyses of smoking behaviors in African Americans. *Transl Psychiatry.* 2012;2:e119.
43. Liu JZ, Tozzi F, Waterworth DM, et al. Meta-analysis and imputation refines the association of 15q25 with smoking quantity. *Nat Genet.* May 2010;42(5):436-440.
44. Amos CI, Wu X, Broderick P, et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet.* Apr 2 2008;40(5):616-622.
45. Hung RJ, McKay JD, Gaborieau V, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature.* Apr 3 2008;452(7187):633-637.
46. Pillai SG, Ge D, Zhu G, et al. A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. *PLoS Genet.* Mar 2009;5(3):e1000421.
47. Cho MH, McDonald ML, Zhou X, et al. Risk loci for chronic obstructive pulmonary disease: a genome-wide association study and meta-analysis. *Lancet Respir Med.* Mar 2014;2(3):214-225.
48. Freathy RM, Ring SM, Shields B, et al. A common genetic variant in the 15q24 nicotinic acetylcholine receptor gene cluster (CHRNA5-CHRNA3-CHRNB4) is associated with a reduced ability of women to quit smoking in pregnancy. *Hum Mol Genet.* Aug 1 2009;18(15):2922-2927.
49. Chen L-S, Baker TB, Piper ME, et al. Interplay of Genetic Risk Factors (*CHRNA5-CHRNA3-CHRNB4*) and Cessation Treatments in Smoking Cessation Success. *American Journal of Psychiatry.* 2012/7/1 2012;169(7):735-742.
50. Bergen AW, Javitz HS, Krasnow R, et al. Nicotinic acetylcholine receptor variation and response to smoking cessation therapies. *Pharmacogenet Genomics.* Feb 2013;23(2):94-103.
51. Kumasaka N, Aoki M, Okada Y, et al. Haplotypes with copy number and single nucleotide polymorphisms in CYP2A6 locus are associated with smoking quantity in a Japanese population. *PLoS One.* 2012;7(9):e44507.
52. Patel YM, Park SL, Han Y, et al. Novel Association of Genetic Markers Affecting CYP2A6 activity and Lung Cancer Risk. *Cancer Res.* Aug 3 2016.
53. Bloom AJ, Harari O, Martinez M, et al. Use of a predictive model derived from in vivo endophenotype measurements to demonstrate associations with a complex locus, CYP2A6. *Hum Mol Genet.* Jul 1 2012;21(13):3050-3062.
54. Lerman C, Jepson C, Wileyto EP, et al. Genetic variation in nicotine metabolism predicts the efficacy of extended-duration transdermal nicotine therapy. *Clin Pharmacol Ther.* May 2010;87(5):553-557.
55. Chen LS, Bloom AJ, Baker TB, et al. Pharmacotherapy effects on smoking cessation vary with nicotine metabolism gene (CYP2A6). *Addiction.* Jan 2014;109(1):128-137.
56. Saccone SF, Hinrichs AL, Saccone NL, et al. Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. *Hum Mol Genet.* Jan 1 2007;16(1):36-49.
57. Zhu AZ, Renner CC, Hatsukami DK, Benowitz NL, Tyndale RF. CHRNA5-A3-B4 genetic variants alter nicotine intake and interact with tobacco use to influence body weight in Alaska Native tobacco users. *Addiction.* Oct 2013;108(10):1818-1828.
58. Bierut LJ, Stitzel JA, Wang JC, et al. Variants in nicotinic receptors and risk for nicotine dependence. *Am J Psychiatry.* Sep 2008;165(9):1163-1171.
59. Amos CI, Gorlov IP, Dong Q, et al. Nicotinic acetylcholine receptor region on chromosome 15q25 and lung cancer risk among African Americans: a case-control study. *J Natl Cancer Inst.* Aug 4 2010;102(15):1199-1205.

60. Wang JC, Grucza R, Cruchaga C, et al. Genetic variation in the CHRNA5 gene affects mRNA levels and is associated with risk for alcohol dependence. *Mol Psychiatry.* Apr 15 2009;14(5):501-510.
61. Wang JC, Cruchaga C, Saccone NL, et al. Risk for nicotine dependence and lung cancer is conferred by mRNA expression levels and amino acid change in CHRNA5. *Hum Mol Genet.* Aug 15 2009;18(16):3125-3135.
62. Smith RM, Alachkar H, Papp AC, et al. Nicotinic alpha5 receptor subunit mRNA expression is associated with distant 5' upstream polymorphisms. *Eur J Hum Genet.* Aug 11 2010;19(1):76-83.
63. Chen LS, Saccone NL, Culverhouse RC, et al. Smoking and genetic risk variation across populations of European, Asian, and African American ancestry--a meta-analysis of chromosome 15q25. *Genet Epidemiol.* May 2012;36(4):340-351.
64. Li MD, Yoon D, Lee JY, et al. Associations of variants in CHRNA5/A3/B4 gene cluster with smoking behaviors in a Korean population. *PLoS One.* 2010;5(8):e12183.
65. Sarginson JE, Killen JD, Lazzeroni LC, et al. Markers in the 15q24 nicotinic receptor subunit gene cluster (CHRNA5-A3-B4) predict severity of nicotine addiction and response to smoking cessation therapy. *Am J Med Genet B Neuropsychiatr Genet.* Apr 2011;156B(3):275-284.
66. Zhu AZ, Zhou Q, Cox LS, et al. Association of CHRNA5-A3-B4 SNP rs2036527 with smoking cessation therapy response in African-American smokers. *Clin Pharmacol Ther.* Aug 2014;96(2):256-265.
67. Munafo MR, Johnstone EC, Walther D, et al. CHRNA3 rs1051730 genotype and short-term smoking cessation. *Nicotine Tob Res.* Oct 2011;13(10):982-988.
68. Chen LS, Baker TB, Jorenby D, et al. Genetic variation (CHRNA5), medication (combination nicotine replacement therapy vs. varenicline), and smoking cessation. *Drug Alcohol Depend.* Sep 1 2015;154:278-282.
69. Kaufman JS, Harper S. Deficiency of the odds ratio for common outcomes. *Am J Psychiatry.* Oct 2012;169(10):1118; author reply 1118-1119.
70. Chen L-S, Baker TB, Bierut LJ. Response to Kaufman and Harper Letter. *Am J Psychiatry.* Oct 2012;169(10):1118-1119.
71. Leung T, Bergen A, Munafo MR, et al. Effect of the rs1051730-rs16969968 variant and smoking cessation treatment: a meta-analysis. *Pharmacogenomics.* 2015;16(7):713-720.
72. Tyndale RF, Zhu AZ, George TP, et al. Lack of Associations of CHRNA5-A3-B4 Genetic Variants with Smoking Cessation Treatment Outcomes in Caucasian Smokers despite Associations with Baseline Smoking. *PLoS One.* 2015;10(5):e0128109.
73. Chen LS, Baker TB, Piper ME, et al. Interplay of genetic risk (CHRNA5) and environmental risk (partner smoking) on cigarette smoking reduction. *Drug Alcohol Depend.* Oct 1 2014;143:36-43.
74. Ware JJ, Aveyard P, Broderick P, et al. The association of rs1051730 genotype on adherence to and consumption of prescribed nicotine replacement therapy dose during a smoking cessation attempt. *Drug Alcohol Depend.* Jun 1 2015;151:236-240.
75. King DP, Paciga S, Pickering E, et al. Smoking cessation pharmacogenetics: analysis of varenicline and bupropion in placebo-controlled clinical trials. *Neuropsychopharmacology.* Feb 2011;37(3):641-650.
76. Benowitz NL. Pharmacology of nicotine: addiction, smoking-induced disease, and therapeutics. *Annual review of pharmacology and toxicology.* 2009;49:57-71.
77. Tanner JA, Chenoweth MJ, Tyndale RF. Pharmacogenetics of nicotine and associated smoking behaviors. *Curr Top Behav Neurosci.* 2015;23:37-86.

78. Nakajima M, Yamamoto T, Nunoya K, et al. Role of human cytochrome P4502A6 in C-oxidation of nicotine. *Drug metabolism and disposition: the biological fate of chemicals.* Nov 1996;24(11):1212-1217.
79. Nakajima M, Yamamoto T, Nunoya K, et al. Characterization of CYP2A6 involved in 3'-hydroxylation of cotinine in human liver microsomes. *The Journal of pharmacology and experimental therapeutics.* May 1996;277(2):1010-1015.
80. Bruick RK, McKnight SL. A conserved family of prolyl-4-hydroxylases that modify HIF. *Science.* Nov 9 2001;294(5545):1337-1340.
81. Bloom AJ, Baker TB, Chen LS, et al. Variants in two adjacent genes, EGLN2 and CYP2A6, influence smoking behavior related to disease risk via different mechanisms. *Hum Mol Genet.* Jan 15 2014;23(2):555-561.
82. Swan GE, Lessov-Schlaggar CN, Bergen AW, et al. Genetic and environmental influences on the ratio of 3'hydroxycotinine to cotinine in plasma and urine. *Pharmacogenet Genomics.* May 2009;19(5):388-398.
83. Chenoweth MJ, Novalen M, Hawk LW, Jr., et al. Known and novel sources of variability in the nicotine metabolite ratio in a large sample of treatment-seeking smokers. *Cancer Epidemiol Biomarkers Prev.* Sep 2014;23(9):1773-1782.
84. Bloom J, Hinrichs AL, Wang JC, et al. The contribution of common CYP2A6 alleles to variation in nicotine metabolism among European-Americans. *Pharmacogenet Genomics.* Jul 2011;21(7):403-416.
85. Lerman C, Tyndale R, Patterson F, et al. Nicotine metabolite ratio predicts efficacy of transdermal nicotine for smoking cessation. *Clin Pharmacol Ther.* Jun 2006;79(6):600-608.
86. Schnoll RA, Patterson F, Wileyto EP, et al. Nicotine metabolic rate predicts successful smoking cessation with transdermal nicotine: a validation study. *Pharmacol Biochem Behav.* Mar 2009;92(1):6-11.
87. Kaufmann A, Hitsman B, Goelz PM, et al. Rate of nicotine metabolism and smoking cessation outcomes in a community-based sample of treatment-seeking smokers. *Addictive behaviors.* Dec 2015;51:93-99.
88. Ho MK, Mwenifumbo JC, Al Koudsi N, et al. Association of nicotine metabolite ratio and CYP2A6 genotype with smoking cessation treatment in African-American light smokers. *Clin Pharmacol Ther.* Jun 2009;85(6):635-643.
89. Piper ME, Smith SS, Schlam TR, et al. A randomized placebo-controlled clinical trial of 5 smoking cessation pharmacotherapies. *Arch Gen Psychiatry.* Nov 2009;66(11):1253-1262.
90. Bird A. Perceptions of epigenetics. *Nature.* May 24 2007;447(7143):396-398.
91. Bibikova M, Barnes B, Tsan C, et al. High density DNA methylation array with single CpG site resolution. *Genomics.* Oct 2011;98(4):288-295.
92. Breitling LP, Yang R, Korn B, Burwinkel B, Brenner H. Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. *Am J Hum Genet.* Apr 8 2011;88(4):450-457.
93. Joubert BR, Haberg SE, Nilsen RM, et al. 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ Health Perspect.* Oct 2012;120(10):1425-1431.
94. Philibert RA, Beach SR, Brody GH. Demethylation of the aryl hydrocarbon receptor repressor as a biomarker for nascent smokers. *Epigenetics.* Nov 2012;7(11):1331-1338.
95. Wan ES, Qiu W, Baccarelli A, et al. Cigarette smoking behaviors and time since quitting are associated with differential DNA methylation across the human genome. *Hum Mol Genet.* Jul 1 2012;21(13):3073-3082.
96. Zeilinger S, Kuhnel B, Klopp N, et al. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS One.* 2013;8(5):e63812.

97. Gao X, Jia M, Zhang Y, Breitling LP, Brenner H. DNA methylation changes of whole blood cells in response to active smoking exposure in adults: a systematic review of DNA methylation studies. *Clin Epigenetics.* 2015;7:113.
98. Shenker NS, Polidoro S, van Veldhoven K, et al. Epigenome-wide association study in the European Prospective Investigation into Cancer and Nutrition (EPIC-Turin) identifies novel genetic loci associated with smoking. *Hum Mol Genet.* Mar 1 2013;22(5):843-851.
99. Dogan MV, Shields B, Cutrona C, et al. The effect of smoking on DNA methylation of peripheral blood mononuclear cells from African American women. *BMC Genomics.* 2014;15:151.
100. Elliott HR, Tillin T, McArdle WL, et al. Differences in smoking associated DNA methylation patterns in South Asians and Europeans. *Clin Epigenetics.* 2014;6(1):4.
101. Tsaprouni LG, Yang TP, Bell J, et al. Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. *Epigenetics.* Oct 2014;9(10):1382-1396.
102. Philibert RA, Beach SR, Lei MK, Brody GH. Changes in DNA methylation at the aryl hydrocarbon receptor repressor may be a new biomarker for smoking. *Clin Epigenetics.* 2013;5(1):19.
103. Markunas CA, Xu Z, Harlid S, et al. Identification of DNA methylation changes in newborns related to maternal smoking during pregnancy. *Environ Health Perspect.* Oct 2014;122(10):1147-1153.
104. Richmond RC, Simpkin AJ, Woodward G, et al. Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Hum Mol Genet.* Apr 15 2015;24(8):2201-2217.
105. Sun YV, Smith AK, Conneely KN, et al. Epigenomic association analysis identifies smoking-related DNA methylation sites in African Americans. *Hum Genet.* Sep 2013;132(9):1027-1037.
106. Zhang Y, Yang R, Burwinkel B, et al. F2RL3 methylation in blood DNA is a strong predictor of mortality. *Int J Epidemiol.* Aug 2014;43(4):1215-1225.
107. Reynolds LM, Wan M, Ding J, et al. DNA Methylation of the Aryl Hydrocarbon Receptor Repressor Associations With Cigarette Smoking and Subclinical Atherosclerosis. *Circ Cardiovasc Genet.* Oct 2015;8(5):707-716.
108. Dietz PM, Homa D, England LJ, et al. Estimates of nondisclosure of cigarette smoking among pregnant and nonpregnant women of reproductive age in the United States. *Am J Epidemiol.* Feb 1 2011;173(3):355-359.
109. Connor Gorber S, Schofield-Hurwitz S, Hardt J, Levasseur G, Tremblay M. The accuracy of self-reported smoking: a systematic review of the relationship between self-reported and cotinine-assessed smoking status. *Nicotine Tob Res.* Jan 2009;11(1):12-24.
110. Benowitz NL, Hukkanen J, Jacob P, 3rd. Nicotine chemistry, metabolism, kinetics and biomarkers. *Handb Exp Pharmacol.* 2009(192):29-60.
111. Sandberg A, Skold CM, Grunewald J, Eklund A, Wheelock AM. Assessing recent smoking status by measuring exhaled carbon monoxide levels. *PLoS One.* 2011;6(12):e28864.
112. Jarvis MJ, Russell MA, Benowitz NL, Feyerabend C. Elimination of cotinine from body fluids: implications for noninvasive measurement of tobacco smoke exposure. *Am J Public Health.* Jun 1988;78(6):696-698.
113. Wall MA, Johnson J, Jacob P, Benowitz NL. Cotinine in the serum, saliva, and urine of nonsmokers, passive smokers, and active smokers. *Am J Public Health.* Jun 1988;78(6):699-701.
114. Relton CL, Hartwig FP, Davey Smith G. From stem cells to the law courts: DNA methylation, the forensic epigenome and the possibility of a biosocial archive. *Int J Epidemiol.* Aug 2015;44(4):1083-1093.

115. Shenker NS, Ueland PM, Polidoro S, et al. DNA methylation as a long-term biomarker of exposure to tobacco smoke. *Epidemiology.* Sep 2013;24(5):712-716.
116. Zhang Y, Yang R, Burwinkel B, Breitling LP, Brenner H. F2RL3 methylation as a biomarker of current and lifetime smoking exposures. *Environ Health Perspect.* Feb 2014;122(2):131-137.
117. Teschendorff AE, Yang Z, Wong A, et al. Correlation of Smoking-Associated DNA Methylation Changes in Buccal Cells With DNA Methylation Changes in Epithelial Cancer. *JAMA Oncol.* Jul 2015;1(4):476-485.
118. Beachy SH, Johnson SG, Olsen S, Berger AC, eds. *Refining Processes for the Co-Development of Genome-Based Therapeutics and Companion Diagnostic Tests: Workshop Summary*. Washington, D.C.: The National Academies Press, Institute of Medicine of the National Academies; 2014. Policy BoHS, ed.
119. David SP, Bergen AW, Munafo M, et al. Genomic analysis to guide choice of treatment for smoking cessation. *Cochrane Database of Systematic Reviews.* 2015.
120. Hendershot T, Pan H, Haines J, et al. Using the PhenX Toolkit to Add Standard Measures to a Study. *Curr Protoc Hum Genet.* 2015;86:1 21 21-17.

Table 1. Glossary

Term	Definition
Allele	A particular DNA sequence or state possible at a genetic locus.
Biomarker	A characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention. A biomarker that is intended to substitute for a clinical endpoint may be referred to as a "surrogate endpoint." [27].
CAP	College of American Pathologists: provides CAP accreditation programs for laboratories.
CLIA	Clinical Laboratory Improvements Amendments: United States federal regulatory standards for clinical laboratory testing. CLIA-certification
CNV	Copy number variant: a genetic variant in which a particular stretch of DNA occurs with variable numbers of copies.
DNA Sequencing	The laboratory and analytical processes to determine the nucleotide sequence of a portion or all of a genome.
Epigenetics	Non-genetic sequence modifications of the DNA including DNA methylation, histone modifications, and miRNAs. Epigenetic processes regulate transcription and expression of genes.
GWAS	Genome-wide association study: a method for identifying genetic variants associated with a disease or quantitative trait by genotyping and analyzing a large number of variants across the genomes of multiple individuals who vary in their disease status or trait values.
QTL	Quantitative Trait Locus. A genetic variant that is associated with differential levels of a quantitative trait. Specific types include expression QTLs (eQTLs) associated with differences in gene expression levels and methylation QTLs (meQTLs) associated with

	differences in gene methylation levels.
SNP	Single nucleotide polymorphism: a genetic variant occurring at a specific nucleotide in the DNA sequence.
SNV	Single nucleotide variant: similar to a SNP, but typically used when the variant may be of lower frequency (rarer).

Table 2. Overview of key genetic associations cited in this review, between *CHRNA5* and *CYP2A6* genotypes and traits relevant for smoking and smoking cessation.

Gene	Associated traits/outcomes
CHRNA5	GWAS-significant (p < 5E-08): Nicotine dependence[38-39] Smoking heaviness / cigarettes-per-day[16, 38-43] Biomarkers of smoking (cotinine levels[29-30], carbon monoxide levels[18]) Lung cancer[38, 44-45] Chronic Obstructive Pulmonary Disease (COPD)[46-47] Other: Difficulty quitting in pregnant smokers[48] Delayed smoking cessation and earlier lung cancer diagnosis in smokers [14] Differential response to pharmacological cessation treatment in some clinical trials[49-50]
CYP2A6	GWAS-significant (p < 5E-08): Smoking heaviness / cigarettes-per-day[16, 40, 51] Biomarker of nicotine metabolism: the Nicotine Metabolism Ratio (NMR)[32-34, 52] Other: Genotype-based biomarker of CYP2A6 enzyme activity [53] Response to extended-duration transdermal nicotine therapy [54] Response to pharmacologic smoking cessation treatment [55]